

버스트 교환 방식 CDMA 시스템에서의 패킷 데이터 서비스를 위한 전송률 스케줄링 기법

정희원 김 미 정*, 최 인 찬**, 강 충 구***, Ramesh R. Rao****

A Transmission Rate Scheduling Scheme for Packet Data Service in the Burst-Switching DS/CDMA System

Meejoung Kim*, In-Chan Choi**, Chung Gu Kang***, and Ramesh R. Rao**** *Regular Members*

요 약

본 논문에서는 패킷 데이터 서비스를 지원하는 DS/CDMA 시스템에서의 버스트 교환 방식을 고려한 전송률 제어 기법으로서, 상향 링크의 대기 패킷 버스트들이 겪는 평균 전송 시간을 최소화하기 위해 각 단말들에게 할당될 최적의 전송률을 결정하는 스케줄링 기법을 제안한다. 본 제안 기법은 회선 교환 방식의 음성 서비스와 더불어 패킷 데이터 서비스가 통합될 때 패킷의 길이에 따라 최적의 전송 그룹을 구성하고, 평균 전송 지연 및 전송 완료 시간이 최소가 되도록 전송률과 전송 순서를 결정한다. 일반적으로 모든 데이터 사용자에게 단일의 전송률을 할당하는 방식 및 단일 서버 모델에서 최적으로 알려져 있는 SPTF (Shortest Processing Time First) 방식과 제안된 방식의 성능을 비교하였다.

ABSTRACT

This paper proposes a new packet rate scheduling scheme for a non-real time data service over the uplink of a burst switching-based direct sequence code division multiple access (DS/CDMA) system to support the integrated voice/data service. We consider the most general form of optimization problem formulation to determine the optimal number of transmission-time groups along with their data rates, which minimize the average packet transmission delay. An ordered packet length-based group-wise transmission (OLGT) scheme is proposed as a simple heuristic solution approach to this problem and present some analytical results for performance comparison with other possible schemes.

1. 서론

제 3세대 이동통신 시스템의 가장 기본적인 설계 목표는 높은 전송률과 더불어 주파수 효율성의 극대화를 실현하는 것이다. 이를 위해 회선 교환 방식의 기존 음성 서비스뿐만 아니라 패킷 데이터 서비스의 통합을 전제로 하고 있으며, 특히 이를 위해 DS/CDMA 시스템에서는 데이터 트래픽의 특성을 고려하여 다중 접속(multiple access)에 의해 발생하

는 간섭을 적절히 제어함으로써 각 서비스가 요구하는 품질을 만족시키면서 채널의 효율성을 극대화할 수 있는 매체접근제어 (medium access control: MAC) 프로토콜의 구현이 요구된다. 특히, 상향 링크의 경우에는 독립적으로 분산된 이동국들에 대한 트래픽 발생 특성을 알 수 없으므로 MAC 프로토콜을 통해 채널의 효율성을 극대화 시키면서 서비스 트래픽에 따른 지연 시간 등의 QoS를 보장할 수 있어야 한다.

* 고려대학교 BK-21 정보기술사업단(kmj@semi.korea.ac.kr) ** 고려대학교 공과대학 산업시스템정보공학과(icho@korea.ac.kr)

*** 고려대학교 정보통신대학 전파통신공학과(ccgknag@korea.ac.kr)

**** Department of Electrical & Computer Engineering University of California at San Diego(rrao@ece.ucsd.edu)

※ 본 논문은 2002년 4월 JCCI 학술회에서 우수논문으로 선정되어 게재 추천된 논문입니다.

한편, 기존의 회선 교환 방식이 갖는 비효율성을 극복하고, 패킷 교환 방식에서 큐잉 지연으로 인해 발생하는 패킷당 과도한 평균 지연 시간 및 복잡도를 동시에 고려하여 CDMA 시스템에서 웹 브라우징이나 WWW 트래픽과 같이 고속 전송을 요하면서 버스티한 특성을 갖는 데이터 서비스의 성능을 향상시키기 위해 버스트 교환(burst switching)이라는 새로운 개념이 제안된 바 있다^[1]. 이는 이미 IS-95C 과 같은 기존의 상용 시스템에서 고려된 바와 같이 전용 제어 채널을 통해 데이터 버스트 단위로 네트워크내의 트래픽 부하에 따른 간섭을 고려하여 상향 트래픽 링크의 전송률을 동적으로 설정하고, 버스트 단위로 무선 자원을 예약함으로써 패킷 전송 지연 시간의 단축과 주파수 효율성의 극대화를 실현하고자 한다. 즉, 버스트 교환 방식에서는 기존의 패킷 교환 방식과 달리 일련의 패킷으로 구성된 버스트 단위로 무선 자원을 할당하고, 버스트가 종료되면 무선 자원을 해제하게 된다. 이 때 패킷 데이터 사용자가 겪는 지연 시간을 최소화하기 위해 미리 설정된 타이머 값을 기준으로 물리 계층과 상위 계층의 무선 자원에 대한 유지와 해제를 2단계의 절차에 따라 결정하게 된다^[1]. 또한, 버스트 교환 방식 DS/CDMA 시스템에서는 각 데이터 단말의 버스트 발생 특성에 따라 간섭 제한 조건 내에서 할당 가능한 데이터 전송률을 결정하기 위한 스케줄링이 이루어져야 하며, 이는 데이터 서비스의 평균 전송 지연시간을 극소화하는 방향에서 수행되어야 한다. 고속 데이터 전송을 필요로 하는 경우 자원의 예약 요구 및 할당 메시지는 dedicated control channel 을 통하여 이루어진다. 특히 cdma2000에서는 supplemental code channel 요구 메시지는 메시지의 길이나 이웃 셀의 파일럿 정보 등을 포함하는 반면 할당 메시지(assignment message)는 지속시간(duration), 시작시간(start time) 및 패킷의 전송률(rate) 등과 같은 스케줄링 파라미터들을 포함한다^[2].

기존 관련 연구는 하향 링크에서의 QoS를 보장하기 위한 전송률 스케줄링(rate scheduling) 문제 [3-5] 및 상향 링크를 고려한 문제들을 다루고 있다^[6-9]. [3]에서는 각 시분할 슬롯마다 한명의 사용자를 서비스하는 단일 서버 모델을 고려하였다. 즉 한 번에 한 사용자에게 허용 가능한 최대의 전송률을 할당하고, 이때 잔여 수명이 최소가 되는 패킷을 우선적으로 전송하는 Earliest Deadline Algorithm (EDA)이 최적의 스케줄링 방식이라고 결론을 내렸다. 한편, CDMA 시스템의 특성에 따라 동시에 여

러 사용자들을 다중화하여 서비스할 수 있는 다중 서버 모델에 기반을 두기도 하다. 예를 들어, [4]에서는 한 사용자에게 할당되고 남은 자원이 있는 한 적절하게 설정된 순위와 사용자간의 간섭을 고려하여 다른 사용자에게 남은 자원을 추가로 할당하는 PEDF (Powered Earliest Deadline First) 방식을 제안하였다. 또한, 유사한 방법으로 [5]에서는 가중치를 둔 잔여 수명을 패킷 길이로 정규화한 동적 우선 순위(dynamic priority)라는 새로운 메트릭(metric)을 고려하였다. 제안된 메트릭은 패킷의 길이가 길고 잔여 수명이 짧은 패킷에 우선 순위를 주는 것이며, 간섭 제약식이 만족하는 한 우선 순위에 따라 패킷들에게 전송률을 할당한다. 기존의 스케줄링 방식은 우선 순위를 가진 패킷에 높은 전송률을 할당하는 것이므로 고속 데이터 서비스에 적합하지만 순위가 낮은 패킷은 서비스를 받지 못할 수 있다는 공정성 문제가 단점이 될 수 있다. 단일 서버 모델에서 이와 같은 공정성 문제를 해결하기 위해 제시된 대표적인 접근 방법은 cdma2000 1x EV-DO 시스템에서 제안된 Proportional Fairness 알고리즘이다.

버스트 교환 방식을 고려한 상향 링크는 시스템 내에서 동시에 다중 버스트 데이터 전송률 할당 요구가 이루어지므로 이에 따른 전송 스케줄링이 이루어져야 한다. [6]에서는 회선 교환 방식하의 DS/CDMA 시스템에서 음성/데이터 통합 서비스에서의 데이터 사용자의 수용을 개선하는 방식을 고려하였고, [7]에서는 최대 전력과 비실시간 사용자들간의 간섭과 최대 가능한 전송 전력의 제한을 고려하면서 채널의 효율성을 극대화하기 위한 전송률 할당 방식을 고려하였다. [8]에서는 전송 지연 시간을 고려한 최적화 정식을 통해 단일 서버 모델과 다중 서버 모델간의 득실관계를 고려한 최적해의 속성을 제시하였으나, 실질적인 구현 가능 정도의 복잡도를 갖지 못한다. 한편, [9]에서는 모든 활성 사용자들이 동시에 서비스되는 다중 서버 모델에서 평균 전송 지연을 최소화하면서 구현이 단순한 그룹 전송 기법을 제안하였다.

본 논문에서는 버스트 교환 방식을 고려한 상향 링크에서 여러 명의 활성화 된 데이터 사용자가 동시에 전송하는 다중 서버 모델을 고려한다. 모든 활성화 된 데이터 사용자는 동시에 스케줄링 되고 각각의 데이터 전송률은 평균 전송 지연이 최소가 되도록 결정된다. 일반적으로 최적의 스케줄링 방식은 간섭 제약식과 평균 전송 지연에 관한 식을 간단하

게 도출하기가 어렵기 때문에 실제 구현에서는 적용하기가 어렵다. 본 논문에서는 최적 스케줄링 방식의 특성을 재고하여 수학적 구조의 틀을 제공하며, 이를 통해 새로운 방식을 도출하고자 한다.

본 논문의 구성은 다음과 같다. II장에서는 시스템 모델 및 기존의 패킷 전송 방식을 설명하고 스케줄링 방식의 최적화에 대한 수학적 모델을 제시한다. III장에서는 새로운 스케줄링 방식을 제안하고 성능을 분석한다. IV장에서는 성능 분석 결과 및 앞으로의 연구 진행 방향을 언급한다.

II. 시스템 모델 및 스케줄링 문제의 정식화

1. 시스템 모델

확산 대역폭이 W 이며, 음성과 데이터의 통합 서비스를 지원하는 단일 셀의 DS/CDMA 시스템에서 상향 링크를 고려한다. 음성 서비스에 대해서는 음성 활성 상태 검출(voice activity detection)이 가능하며, R_v (bps)의 고정된 전송률을 할당하는 회선 교환 방식이 적용된다. 반면, 데이터 서비스의 경우에는 시스템 내의 사용자 수에 의해 야기되는 간섭의 정도에 따라 데이터 전송률 R_d (bps)를 패킷 단위로 동적으로 할당한다. 새로운 패킷 버스트가 발생하면 주기적으로 반복되는 스케줄링 시점까지 대기하고, 그 시점에서 모든 단말들의 패킷 전송 요구 사항에 따라 신호 대 간섭 비의 제약 조건을 만족하는 범위 내에서 최대 허용 가능한 전송률을 할당 받게 된다. 일단 데이터 전송률이 결정되면 해당 전용 트래픽 채널이 각 사용자에게 예약되고, 전송이 종료되거나 새로운 스케줄링이 시작되는 시점까지 고정적으로 할당된다. 버스트의 길이는 특정 크기에 의해 제한되고, 스케줄링 주기는 최대 버스트의 길이보다 길다고 가정하며, 또한 버스트가 시스템으로 입력될 때 버스트 수락제어(Burst Admission Control: BAC)가 수행되어 스케줄링 시점까지 대기하는 큐잉 지연은 충분히 작다고 가정한다. 다음에서 고려할 스케줄링 알고리즘은 전송 지연 시간의 최적화에 초점을 맞춘다.

시스템 내에 있는 음성과 데이터 사용자의 수는 각각 N_v 과 N_d 로 나타낸다. 음성과 데이터 서비스의 요구 품질에 따라서 필요한 비트 오류율이 있으며, 이를 만족하기 위해 요구되는 비트 에너지 대 잡음비 (E_b/N_0)가 각각 η_v 와 η_d 로 주어졌다고 하자. 열잡음을 제외한 다른 잡음은 없다고 가정하고, 전력제어가 완벽하다고 가정하면 음성 및 데이터의

통합 서비스를 위해 다음과 같은 정규화 된 제약식을 만족해야 한다^[10].

$$\frac{N_v}{\frac{W}{\alpha_v R_v \eta_v} + 1} + \frac{N_d}{\frac{W}{\alpha_d R_d \eta_d} + 1} \leq 1 \quad (1)$$

(1) 식에서 α_v 는 회선 교환 방식에 따른 음성 트래픽의 활성비(voice activity factor)를 나타내며, 데이터 트래픽 채널은 버스트 전송이 끝날 때까지 채널을 고정적으로 사용하므로 데이터 서비스의 활성비는 1로 가정한다($\alpha_v = 1$). (1)식에서 음성과 데이터 사용자에게 할당되는 전력비를 각각 Δ_v 와 Δ_d 로 표현한다.

2. 스케줄링 문제의 정식화

본 논문에서는 전송 지연 시간을 최소화하기 위한 스케줄링 문제를 다루고자 한다. 일반적인 전송 방식으로는 패킷 전송을 요청하는 모든 사용자들이 동시에 전송하도록 하는 방식과 한번에 한 사용자씩 순차적으로 전송하도록 하는 방식으로 크게 구분할 수 있다. 후자의 경우에 대하여는 처리시간이 가장 짧은 사용자부터 순차적으로 순서를 정해 전송하는 Shortest Processing Time first (SPTF) 방식이 최적의 스케줄링 방식으로 알려져 있다^[11]. 그러나, 위의 전송 방식에서 단말의 최대 송출 가능 전력의 한계 또는 채널의 상태에 따른 적응 변조 방식 등에 의해 결정되는 기타 물리계층의 제약에 따라 최대 허용 가능한 전송률이 R_M (bps)로 제한될 수 있다. 반면에 전자의 경우 중에 패킷 전송 지연 시간을 최소화하기 위한 가장 단순한 방법은 모든 사용자들에게 단일의 전송률을 할당하되, (1)의 제약식을 만족하면서 패킷 전송 지연시간을 최소화하기 위한 전송률을 결정하는 단일 전송률 (Non-Grouping Transmission: NGT) 방식이다^[6].

본 논문에서는 위에서 언급한 두 가지 스케줄링 방식을 혼합하여 특정 순서에 의해 동시에 전송되는 일부 사용자들의 그룹화와 각 그룹에서 전송되는 각 사용자들의 개별 전송률을 결정하는 방식을 고려하고자 한다. 그룹을 나누는 것은 QoS로 특징 지워진 우선 순위에 따른 순위 메트릭을 따른다. [5]에서는 가중치를 준 잔여 수명을 패킷 길이로 정규화한 pseudoprobability라는 메트릭을 사용하는데 본 논문에서는 분석의 단순화를 위하여 패킷 길이만을 고려한 메트릭을 사용한다.

우선 순위에 따른 전송 순서에 따라 정해진 스케

줄링 그룹이 정해지고, 각 그룹에서 동시에 전송되는 사용자의 수를 N 이라고 하자. 이 방식의 두 가지 특수한 경우로서 패킷 전송을 요청하는 모든 사용자에게 대해 동시에 자원을 할당하는 경우와 STPF 방식의 경우가 포함된다. 즉, 전자의 경우는 $N=N_d$, 후자의 경우는 $N=1$ 에 해당되며, 제안된 혼합 방식인 경우는 $1 < N < N_d$ 가 된다. 이 경우에 i 번째 데이터 사용자의 패킷 길이를 L_i 라 하고 그들이 할당 받는 데이터 전송률을 $R_i^{(d)}$, $i=1, \dots, N$,라 하면 스케줄링 문제를 다음과 같이 구별되는 두 가지의 최적화 정식으로 표현할 수 있다.

$$(P0) \quad \min_{\mathbf{R} \in F_0} \frac{1}{N} \sum_{i=1}^N \frac{L_i}{R_i^{(d)}}$$

$$(P1) \quad \min_{\mathbf{R}} \max_{1 \leq i \leq N} \left(\frac{L_i}{R_i^{(d)}} \right)$$

(P0)와 (P1)은 각각 평균 전송 지연의 최소화 관점과 전송 완료 시간을 최소화하는 관점에서의 최적화 정식이며 feasible set F_0 는 다음과 같이 주어진다.

$$F_0 = \{ \mathbf{R}^{(d)} = (R_1^{(d)}, R_2^{(d)}, \dots, R_N^{(d)}) : 0 < R_i^{(d)}, \Delta_v + \sum_{i=1}^N P_i \leq 1,$$

$$P_i = 1/(W/(R_i^{(d)}\eta_d) + 1), i=1, \dots, N \}$$

여기서 P_i 는 정규화 된 전력이며 R_M 은 고려하지 않는다. 정식 (P0)와 (P1)에 대하여 다음과 같은 Proposition들을 얻는다.

Proposition 1. (P0)의 최적해 $(\tilde{R}_1^{(d)}, \tilde{R}_2^{(d)}, \dots, \tilde{R}_N^{(d)})$ 는 다음 두 식을 만족하는 벡터 $\mathbf{P} = (P_1, \dots, P_N)$ 에 의하여 얻는다.

$$\Delta_v + \sum_{i=1}^N P_i = 1, \quad \frac{L_1}{P_1^2} = \frac{L_2}{P_2^2} = \dots = \frac{L_N}{P_N^2} \quad (2)$$

(P0)를 최소화하는 전송률 $\tilde{R}_i^{(d)}$ 과 그 때의 (P0)의 값 $E_{lb}(D)$ 는 다음과 같다.

$$\tilde{R}_i^{(d)} = \frac{(1-\Delta_v)W\sqrt{L_i}}{\eta_d \left(\sum_{i=1}^N \sqrt{L_i} - (1-\Delta_v)\sqrt{L_i} \right)}$$

$$E_{lb}(D) = \frac{\eta_d}{NW} \left(\frac{\left(\sum_{i=1}^N \sqrt{L_i} \right)^2}{1-\Delta_v} - \sum_{i=1}^N L_i \right)$$

Proposition 2. (P1)의 최적해 $(\bar{R}_1^{(d)}, \bar{R}_2^{(d)}, \dots, \bar{R}_N^{(d)})$ 는 다음을 두 식을 만족하는 벡터 $\mathbf{P} = (P_1, \dots, P_N)$ 에 의하여 얻는다.

$$\Delta_v + \sum_{i=1}^N P_i = 1, \quad \frac{L_1}{\bar{R}_1^{(d)}} = \frac{L_2}{\bar{R}_2^{(d)}} = \dots = \frac{L_N}{\bar{R}_N^{(d)}} \quad (3)$$

Proposition 1,2는 주어진 자원(전력)을 최대한 사용하는 것이 최적의 필요 조건임을 의미하며 Proposition 1은 (2)의 등식이 만족되도록 전송률을 할당하고, Proposition 2는 모든 사용자가 동시에 전송이 완료되도록 전송률을 할당하는 것이 (P0)와 (P1)을 각각 최적화하는 것임을 나타낸다. Proposition 1에 대한 증명은 본 논문의 부록 1에서 별도로 제시하였고, Proposition 2는 N 차 방정식을 풀어야 하므로 일반적인 형태의 해는 없지만 (3)식의 증명은 부록 2에서 볼 수 있다. (P0)의 값은 동시에 전송하는 스케줄링 기법에서의 최적의 평균 전송 지연이므로 평균 전송 지연의 *lower bound* 라고 하자.

가장 일반적인 형태의 최적의 스케줄링 문제는 전력 제한식을 만족하면서 모든 패킷을 전송하는데 필요한 평균 전송 시간을 최소화하기 위한 각 패킷의 전송 시간과 해당 전송률을 결정하는 것이다. 특정 순위 메트릭에 의하여 결정되는 순서에 의하여 전송되는 경우를 모두 포함하는 일반적인 형태의 최적화 정식 (Q0)와 (Q1)을 다음과 같이 제안한다.

$$(Q0) \quad \min_{(\mathbf{t}, \mathbf{R}) \in F_1} \max_{1 \leq i \leq N_d} \left(t_i + \frac{L_i}{R_i} \right)$$

여기서 L_i 는 순위 메트릭에 의하여 결정된 순서에 대응하는 패킷의 길이라고 하고(본 논문에서는 순위 메트릭으로서 실제의 패킷 길이를 고려함), 일반성을 잃지 않고 $t_1 \leq \dots \leq t_{N_d}$ 를 가정하며 feasible set F_1 은 다음과 같이 주어진다.

$$F_1 = \{ (\mathbf{t}, \mathbf{R}) = ((t_1, \dots, t_{N_d}), (R_1, \dots, R_{N_d})) : 0 < R_i \leq R_M, \sum_{i=1}^{n_t} P_i \leq 1 - \Delta_v \}$$

여기서, n_t 는 시간 t 에 동시에 전송하고 있는 패킷의 수이고, P_i 는 그 때의 i 번째 패킷에 할당되는 전력으로써 $P_i = 1/(W/(R_i\eta_d) + 1)$ 로 주어진다. 정식 (Q0)는 패킷의 길이 순서대로 전송되는 것이 아니므로 각각의 t_i 들은 상황에 따른 변수이다. 그러므

로 일반적인 해법을 찾기 어려운 문제이므로 (NP-hard) 휴리스틱한 접근 방법을 통해 보다 단순한 해법을 제시하고자 한다. 여기서 고려하는 해법은 전송할 패킷들을 다수의 전송 시간 그룹 (transmission-time group)으로 구분하고, 각 전송 시간 그룹은 동일한 수의 패킷을 포함한다. 여기서 전송 시간 그룹은 전송의 순서를 의미하는 것으로 앞의 전송 시간 그룹에 속한 패킷의 전송의 시작은 뒤의 전송 시간 그룹에 속한 패킷의 전송 시작에 우선하여 이루어져야 한다는 것이다. 이와 같이 전송 시간 그룹을 고려한 최적화는 다음과 같은 문제로 볼 수 있다.

$$(Q1) \quad \min_{(t, R, n, r)} \max_{1 \leq j \leq n+1} (t_{ji} + \frac{L_{ji}}{R_{ji}})$$

여기서 t_{ji} 는 스케줄링이 시작되는 시점을 기준으로 실제로 j 번째 그룹에 속한 길이가 L_{ji} 인 i 번째 패킷의 전송이 허락되는 시점까지의 시간이며 feasible set F_2 는 다음과 같다.

$$F_2 = \{(t, R, n, r) = ((t_{11}, \dots, t_{1N}), (t_{21}, \dots, t_{2N}), \dots, (t_{n+1,1}, \dots, t_{n+1,N}), (R_{11}, \dots, R_{1N}), \dots, (R_{n+1,1}, \dots, R_{n+1,N}), (N, \dots, N, r)) : t_{ji} \geq 0, t_{ji} \leq t_{j+1,i}, t_{ji} \leq t_{j+1,i}, 0 < R_{ji} \leq R_M, \mathbf{P}e \leq (1 - \Delta, \mathbf{e}, nN + r = N_d, j = 1, \dots, n, i = 1, \dots, N, r = 0, \dots, N - 1)\}$$

여기서 \mathbf{e} 는 모든 성분이 1로 이루어진 $N \times 1$ 열벡터이고, 행렬 \mathbf{P} 는 전력 제한식 $\Delta_r + \sum_{i=1}^{n_i} P_i \leq 1, P_i = 1/(W/R_i \eta_d + 1)$, 을 만족하는 것이며, n_i 는 시간 t 에 동시에 전송하고 있는 패킷의 수이고, P_i 는 그 때의 i 번째 패킷에 할당되는 전력으로서 다음과 같이 주어진다.

$$(\mathbf{P})_{i,r} = \begin{cases} P_{i,r}, & i = 1, \dots, N \\ 0, & \text{otherwise} \end{cases}$$

여기서 \bullet 는 t 시간에 전송하는 패킷이 속한 전송 시간 그룹을 나타내며 전송을 요청한 단말이 N_d 개 있는 경우에 전송 시간 그룹은 r 에 따라, n 개 또는 $n+1$ 개의 전송 시간 그룹으로 나뉜다. 각 전송 시간 그룹은 Proposition 1을 만족하도록 전력을 할당하여 전송한다. (Q0)와 (Q1)이 동일(equivalent)하다는 것은 (Q0)와 (Q1)을 최소화하는 벡터들을 일대일 대응 관계를 설정함으로써 직관적으로 명백한 사실이나 부록 3에서 증명을 볼 수 있다.

Proposition 1과2에서 언급된 바와 같이 최적의

스케줄링 방식은 전송 시간을 최소화하면서 동일한 시간에 패킷의 전송이 완료되어야 하지만 실제로 구현이 어려우므로 다음과 같은 새로운 최적화 문제를 고려하며, 휴리스틱한 방법으로 접근하고자 한다.

$$(OPT) \quad \min_{(t, R, n) \in F_2} \frac{1}{n+1} \sum_{j=1}^{n+1} \max_{1 \leq i \leq N} (t_{ji} + \frac{L_{ji}}{R_{ji}})$$

최적화 정식 (OPT)는 정식 (P0)와 (Q1)을 결합한 형태로 총 전송 시간을 최소화하는 문제를 여러 개의 연속적인 전송 시간 그룹에서의 평균 전송 지연을 최소화하는 문제로 나누어서 보는, 즉 각 전송 시간 그룹에 최적의 데이터 전송률을 할당함으로써 평균 전송 지연을 최소화하고 모든 패킷의 총 전송 시간을 최소화하고자 하는 것이다.

III. 스케줄링 방식의 제안 및 성능 분석

본 장에서는 II장에서 언급한 기존의 스케줄링 방식의 단점을 보완하면서, 동시에 구현이 단순하고 시스템과 사용자 관점 모두에서의 최적화 방식 (OPT)와 유사한 성능을 나타낼 수 있는 스케줄링 방식을 제안한다. 본 논문에서는 전송 우선 순위를 패킷의 길이만으로 결정하였지만 우선 순위를 결정하는 메트릭을 여러 가지 파라미터, 예를 들면 시스템 내에 들어와서 전송이 시작될 때까지의 대기 시간, 지연 제한 시간, 패킷의 길이, 단말과 기지국과의 거리 등, 를 고려한 스케줄링 문제를 다루는 경우에도 유용한 알고리즘이 될 수 있다. 본 논문에서는 정식 (OPT)의 휴리스틱한 접근 방식으로써 Ordered Length-based Group-wise Transmission (OLGT) 방식을 제안한다.

OLGT 방식은 데이터 사용자를 순위 메트릭에 따라 여러 개의 전송 그룹으로 분류하고 각 그룹은 LGT([9] 참조)와 같은 방식으로 패킷의 길이에 따라 두개의 서브 그룹으로 나누어 전송률을 할당하는 방식, 즉 각 전송 그룹에 LGT 방식을 적용하는 접근이다.

데이터 사용자의 패킷 길이 $\{L_i\}_{i=1}^{N_d}$ 는 서로 독립이고 평균 길이가 $1/\mu$ 인 지수분포를 따른다고 가정한다. (independent identically distributed: i.i.d.). $\{L_i\}$ 들이 연속인 분포 함수를 따르므로 일반성을 잃지 않고 패킷의 길이는 모두 다르다고 가정한다. 다음과 같이 패킷 길이를 증가하는 순서로 재배열한 순서

통계량(order statistics) $\{L_{(i)}\}$ 를 고려하자.

$$L_{(1)} < L_{(2)} < \dots < L_{(k)} < L_{(k+1)} < \dots < L_{(N_d)}$$

확률변수 $L_{(1)}, L_{(2)} - L_{(1)}, \dots, L_{(N_d)} - L_{(N_d-1)}$ 들은 서로 독립이며, 다음과 같은 확률 밀도 함수를 갖는다^[12].

$$P(L_{(k+1)} - L_{(k)} = t) = (N_d - k)\mu e^{-(N_d - k)\mu}$$

1. 스케줄링 방식의 제안

LGT 방식은 활성화 된 모든 데이터 사용자들 두개의 서브 그룹(전송률 그룹)으로 나누어 각각의 전송률 그룹에 전송 지연이 최소가 되도록 전송률(dual-rate)을 할당하여 동시에 전송하는 방식인 반면, OLGT 방식은 데이터 사용자를 순위 메트릭에 따라 여러 개의 전송 시간 그룹으로 나누어서 한 개의 그룹씩 차례로 전송하는 방식이다. 각 전송 시간 그룹은 LGT 방식에 따라 다시 두 개의 서브 그룹, 즉 전송률 그룹으로 나누어진다. OLGT 방식의 목적은 전송 시간 그룹을 결정하고 각 시간 그룹 내에서 총 전송 완료 시간 및 전송 지연이 최소가 되도록 하는 것이다. 각 전송 시간 그룹은 큐잉 지연을 최소화하기 위하여 특정한 순서로 전송되어야 한다. OLGT의 자세한 설명은 다음과 같다.

전송 제한률을 고려하지 않는 $R_i < \infty$ 의 경우에 우선 두개의 서로 다른 전송 시간 그룹을 고려하고 여러 개의 전송 시간 그룹에 대하여는 유사하게 일반화한다. 기호의 편의상 $k_{i,\min}$ 와 $k_{i,\max}$, $i=1,2$ 를 다음과 같이 정의한다.

$$k_{1,\min} = \lceil b \lceil N_d / 2 \rceil \rceil, \quad k_{1,\max} = \lceil N_d / 2 \rceil$$

$$k_{2,\min} = k_{1,\max} + \lceil b(N_d - \lceil N_d / 2 \rceil) \rceil, \quad k_{2,\max} = N_d$$

여기서 $b = k^* / N_d \sim 0.54$ 는 LGT 방식에서 구한 평균 전송 지연을 최소화하는 전송률 그룹을 나누는 최적값이다. 서로 다른 두개의 전송 시간 그룹 $G_1 = \{L_{(1)}, \dots, L_{(k_{1,\max})}\}$, $G_2 = \{L_{(k_{1,\max}+1)}, \dots, L_{(k_{2,\max})}\}$ 에 대하여 LGT 방식의 알고리즘을 이용하여 다음과 같은 두개의 전송률 그룹 $G_{i,\min}, G_{i,\max}$, $i=1,2$,을 구분한다.

$$G_{1,\min} : L_{(1)} < \dots < L_{(k_{1,\min})}, \quad G_{1,\max} : L_{(k_{1,\min}+1)} < \dots < L_{(k_{1,\max})}$$

$$G_{2,\min} : L_{(k_{1,\max}+1)} < \dots < L_{(k_{2,\min})}$$

$$G_{2,\max} : L_{(k_{2,\min}+1)} < \dots < L_{(k_{2,\max})}$$

패킷이 $G_{i,\max}$ 나 $G_{i,\min}$ 에 속하는 사용자에 대해 각각

$R_{\max}^{(OLGT)}$ 과 $R_{\min}^{(OLGT)}$ 의 데이터 전송률을 할당한다. 이 때, N_v 명의 음성 사용자, $G_{1,\min}$ 에 해당하는 패킷을 갖는 $k_{1,\min}$ 명의 데이터 사용자, $G_{1,\max}$ 에 해당하는 패킷을 갖는 $k_{1,\max} - k_{1,\min}$ 명의 데이터 사용자에 할당되는 전력비를 각각 Δ_v , Δ_{\min} , Δ_{\max} 라고 하면 이들은 식 $\Delta_v + \Delta_{\min} + \Delta_{\max} = 1$ 을 만족하며 Δ_{\min} 과 Δ_{\max} 은 각각 다음과 같이 주어진다.

$$\Delta_{\min} = \frac{k_{1,\min}}{W/R_{\min}^{(OLGT)}\eta_d + 1}, \quad \Delta_{\max} = \frac{k_{1,\max} - k_{1,\min}}{W/R_{\max}^{(OLGT)}\eta_d + 1}$$

또한 $G_{1,\max}$ 에 속하는 패킷을 갖는 데이터 사용자에 할당되는 데이터 전송률 $R_{\max}^{(OLGT)}$ 은 다음과 같이 구할 수 있다.

$$R_{\max}^{(OLGT)} = \frac{W}{\eta_d} \cdot \frac{1}{(k_{1,\max} - k_{1,\min}) / (1 - \Delta_v - \Delta_{\min}) - 1}$$

G_1 그룹에 속한 패킷들을 전송한 후 각각의 패킷들의 전송이 끝나는 대로 G_2 그룹에 속한 패킷들을 전송한다. 전송률 $R_{\min}^{(OLGT)}$ ($R_{\max}^{(OLGT)}$)를 할당 받은 $G_{1,\min}$ ($G_{1,\max}$) 전송률 그룹에 속한 패킷들의 전송이 끝나는 대로 같은 전송률을 $G_{2,\min}$ ($G_{2,\max}$) 전송률 그룹에 속한 패킷들에 할당하여 전송한다.

이제 각 전송 시간 그룹에서 전송하는 순서를 결정하는 스케줄링 방식을 고려한다. 전송 완료 시간을 최소화하기 위한 최적의 방식은 Proposition 2에 따르면 각 전송률 그룹의 전송 완료 시간을 동일하게 만드는 것이므로 이것과 SPTF 방식을 절충하여 다음과 같은 전송 순서를 고려한다. 모든 패킷들이 길이가 증가하는 순서로 정렬되면 첫번째 전송 시간 그룹은 $k_{1,\max}$ 개의 짧은 패킷들로 이루어져 있는데 이 그룹은 다시 두 개의 전송률 그룹으로 나뉜다. 전송률 그룹 내의 전송은 정렬된 순서대로, 즉 패킷의 길이가 짧은 순서대로 전송된다. 반면에 $k_{2,\max} - k_{1,\max}$ 개의 패킷들로 이루어진 두 번째 전송 시간 그룹내의 두 개의 전송률 그룹은 각각 패킷의 길이가 긴 순서대로 전송된다. 이 전송 순서는 최대 전송 완료 시간을 최소화 해 주는 방식이다. LGT 방식과 더불어 이 간단한 알고리즘이 정적(OPT)의 휴리스틱한 해를 제공하는 기본이 된다. 이 알고리즘은 전송 시간 그룹이 $n > 2$ 인 일반적인 경우에도 확장되어 적용된다. 즉, 홀수번째 전송 시간 그룹은

각 전송률 그룹에서 패킷의 길이가 짧은 것부터, 짝수번째 전송 시간 그룹은 각 전송률 그룹에서 패킷의 길이가 긴 것부터 그 전 전송 시간 그룹에 속한 패킷들의 전송이 끝나는 순서대로 LGT 방식에 의해 결정된 전송률을 할당하여 전송한다.

그림 1은 $N_d = 6n$ 인 경우, 즉 전송 시간 그룹이 n 개이고 각 전송 시간 그룹은 6개의 패킷을 포함하고 있는 경우의 전송 시간 그룹 G_i , 전송률 그룹 $G_{i,\min}$ 와 $G_{i,\max}$, 그리고 $k_{i,\min}$ 와 $k_{i,\max}$ 를 표시한 것이다. 한편, 그림 2는 그림 1에서 $n = 3$ 인 경우에 OLGT 방식의 패킷 전송 순서를 도시한 것이다.

다음에서는 전송률이 $R_i \leq R_M$ 로 제한되는 경우를 살펴본다. 이때는 가능하면 적은 수의 데이터 사용자에게 전송률을 할당함으로써 간섭을 줄일 수 있으므로 데이터 사용자가 충분히 적은 경우에는 모든 사용자에게 R_M 을 할당할 수 있으면 할당하고 데이터 사용자의 수가 많은 경우에는 동시에 $\lceil bN_{OLGT} \rceil$ 명의 데이터 사용자에게 전송률 $R_{\max}^{(OLGT)} = R_M$ 을 $N_{OLGT} - \lceil bN_{OLGT} \rceil$ 명의 데이터 사용자에게 전송률 $R_{\min}^{(OLGT)} = aR_M$ 을 할당할 때 지연이 최소화 될 것이다. 여기서 N_{OLGT} 는 한 전송 시간 그룹내의 패킷의 수, 즉 최적화 정식 (Q1)에서 동시에 전송하는 패킷의 수 N 으로 다음 식을 만족하는 값이며

$$\Delta_v + \frac{bN_{OLGT}}{W} + 1 + \frac{(1-b)N_{OLGT}}{R_M \eta_d} = 1$$

다음과 같이 표현된다.

$$N_{OLGT} = \left\lfloor \frac{(1-\Delta_v)(W/aR_M \eta_d + 1)(W/R_M \eta_d + 1)}{\{ab + (1-b)W/aR_M \eta_d + 1\}} \right\rfloor$$

여기서 $a = 0.44$ 와 $b = 0.54$ 로서 이는 LGT 방식의 결과에 의한 것이다^[9]. 전송 시간 그룹의 개수는 $n = \lceil N_d / N_{OLGT} \rceil$ 가 된다.

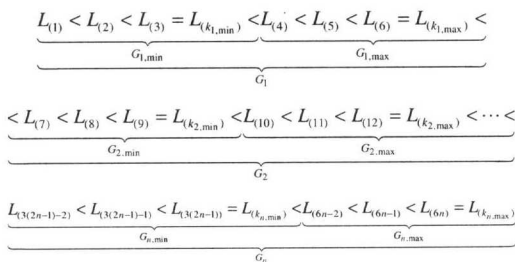


그림 1. 전송 시간그룹 및 전송률 그룹

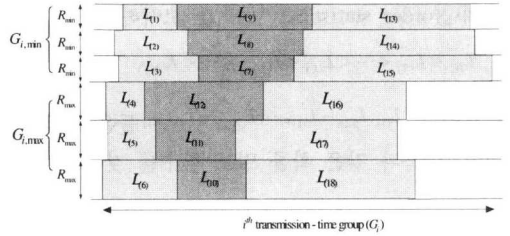


그림 2. OLGT 스케줄링 방식

2. 성능 분석

Case 1. $R_i < \infty$

전송 시간 그룹을 두 개로 하면 G_i 그룹에 속한 패킷들의 전송 지연 시간 $T_i, i = 1, 2$, 는 다음과 같다.

$$T_1 = \sum_{j=1}^{k_{i,\min}} \frac{L_{(j)}}{R_{\min}^{(OLGT)}} + \sum_{j=1}^{k_{i,\max} - k_{i,\min}} \frac{L_{(k_{i,\min} + j)}}{R_{\max}^{(OLGT)}}$$

$$T_2 = \sum_{j=1}^{k_{1,\min}} \left(\frac{L_{(j)}}{R_{\min}^{(OLGT)}} + \frac{L_{(k_{2,\min} - j + 1)}}{R_{\min}^{(OLGT)}} \right) + \sum_{j=1}^{k_{1,\max} - k_{1,\min}} \left(\frac{L_{(k_{1,\min} + j)}}{R_{\max}^{(OLGT)}} + \frac{L_{(k_{2,\max} - j + 1)}}{R_{\max}^{(OLGT)}} \right)$$

G_2 전송 시간 그룹에 속한 각각의 패킷은 G_1 전송 시간 그룹에 속한 한 패킷의 전송 시간만큼 지연이 되므로 T_2 는 G_1 그룹의 패킷의 전송 시간, 즉 큐잉 지연 시간을 포함한다. 그러므로 평균 전송 지연 $E_{OLGT}(D)$ 은 다음과 같이 계산된다.

$$E_{OLGT}(D) = \frac{1}{N_d \mu} \left[\frac{1}{R_{\min}^{(OLGT)}} \sum_{j=1}^{k_{1,\min}} \left\{ 2 \sum_{i=0}^{j-1} \frac{1}{N_d - i} + \sum_{i=0}^{k_{2,\min} - j} \frac{1}{N_d - i} \right\} + \frac{1}{R_{\max}^{(OLGT)}} \sum_{j=1}^{k_{1,\max} - k_{1,\min}} \left\{ 2 \sum_{i=0}^{k_{1,\min} + j - 1} \frac{1}{N_d - i} + \sum_{i=0}^{k_{2,\max} - j} \frac{1}{N_d - i} \right\} \right]$$

일반적으로 n 개의 전송 시간 그룹을 고려하는 경우 $n = \lceil N_d / k_{1,\max} \rceil$ 가 되며, j 번째 전송 시간 그룹에 속한 전송률 그룹의 전송 지연 시간을 각각 $T_{j,\min}, T_{j,\max}$ 라 하면 다음과 같은 recursive 한 식을 만족한다.

$$T_{1,\min} = \sum_{k=1}^{k_{1,\min}} \frac{L_{(k)}}{R_{\min}^{(OLGT)}}$$

$$T_{j,\min} = \sum_{k=j-1,\max+1}^{k_{j,\min}} \left(T_{k_{j,\min} - k + k_{j-2,\max} + 1, \min} + \frac{L_{(k)}}{R_{\min}^{(OLGT)}} \right), \quad 2 \leq j \leq n,$$

$$T_{1,\max} = \sum_{k=k_{1,\min}+1}^{k_{1,\max}} \frac{L_{(k)}}{R_{\max}^{(OLGT)}}$$

$$T_{j,\max} = \sum_{k=k_{j-1,\min}+1}^{k_{j,\max}} \left(T_{k_{j+1,\min} - k + k_{j-2,\max} + 1, \max} + \frac{L_{(k)}}{R_{\max}^{(OLGT)}} \right), \quad 2 \leq j \leq n,$$

여기서 $k_{0,max} = 0$ 이고 $k_{j,min}$ 와 $k_{j,max}$ 는 j 번째 전송 시간 그룹까지의 데이터의 수로서 각각 그 전송 시간 그룹의 첫번째와 두번째 전송률 그룹까지의 전송해야 할 데이터의 수이며 평균 전송 지연은 다음과 같이 계산된다.

$$E_{OLGT}(D) = \frac{1}{N_d \mu} \left[\frac{1}{R_{min}^{OLGT}} \sum_{j=1}^{k_{1,min}} \left\{ n \sum_{i=0}^{j-1} \frac{1}{N_d - i} + (n-1) \sum_{i=0}^{k_{1,min}-j} \frac{1}{N_d - i} \right. \right. \\ \left. \left. + (n-2) \sum_{i=0}^{k_{1,max}+j-1} \frac{1}{N_d - i} + \dots + \sum_{i=0}^{k_{1,max}-j} \frac{1}{N_d - i} \right\} + \frac{1}{R_{max}^{OLGT}} \sum_{j=1}^{k_{1,max}-k_{1,min}} \left\{ n \sum_{i=0}^{k_{1,max}+j-1} \frac{1}{N_d - i} \right. \right. \\ \left. \left. + (n-1) \sum_{i=0}^{k_{2,max}-j} \frac{1}{N_d - i} + (n-2) \sum_{i=0}^{k_{2,min}+j-1} \frac{1}{N_d - i} + \dots + \sum_{i=0}^{k_{2,max}-j} \frac{1}{N_d - i} \right\} \right]$$

Case 2. $R_i \leq R_M$

Case 1과 유사하게 다음과 같이 평균 전송 지연이 주어진다.

$$E_{OLGT}(D) = \frac{1}{N_d \mu} \left[\frac{1}{aR_M} \sum_{j=1}^{k_{1,min}} \left\{ n \sum_{i=0}^{j-1} \frac{1}{N_d - i} + (n-1) \sum_{i=0}^{k_{1,min}-j} \frac{1}{N_d - i} \right. \right. \\ \left. \left. + (n-2) \sum_{i=0}^{k_{1,max}+j-1} \frac{1}{N_d - i} + \dots + \sum_{i=0}^{k_{1,max}-j} \frac{1}{N_d - i} \right\} + \frac{1}{R_M} \sum_{j=1}^{k_{1,max}-k_{1,min}} \left\{ n \sum_{i=0}^{k_{1,max}+j-1} \frac{1}{N_d - i} \right. \right. \\ \left. \left. + (n-1) \sum_{i=0}^{k_{2,max}-j} \frac{1}{N_d - i} + (n-2) \sum_{i=0}^{k_{2,min}+j-1} \frac{1}{N_d - i} + \dots + \sum_{i=0}^{k_{2,max}-j} \frac{1}{N_d - i} \right\} \right]$$

IV. 성능 분석

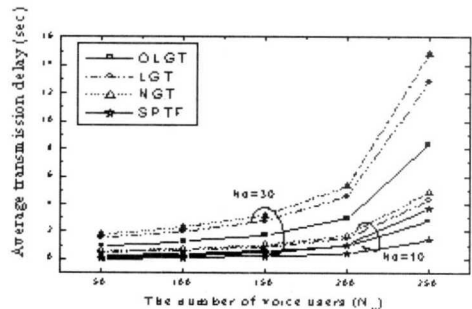
본 분석에서는 다음과 같은 시스템 파라미터를 사용하였다: $W = 5\text{MHz}$, $\alpha_v = 3/8$, $R_v = 9.6\text{kbps}$, $\eta_v = 7\text{dB}$, $\eta_d = 10\text{dB}$, and $1/\mu = 25,000$ bits. 그림 3은 전송 제한률을 고려하지 않은 경우에 SPTF, NGT, LGT, 두개의 전송 시간 그룹을 고려한OLGT의 4가지 다른 전송률 스케줄링 방식의 평균 전송 지연을 비교한 것이다. 이 경우에는 단일 서버 모델이 가장 작은 간섭을 야기하므로 SPTF 방식이 가장 우수한 성능을 나타낼 것임은 명백하다. 음성 사용자에게 관계없이 NGT 방식보다는 제안한 LGT 방식과 OLGT 방식의 성능이 우수하며, 또한 LGT 방식보다는 OLGT 방식의 성능이 훨씬 우수함을 알 수 있다. LGT 방식의 경우 NGT 방식과 비교하여 대략 11%~14%의 성능 향상이 있었으며 OLGT 방식은 대략 60%~70%의 성능 향상이 있음을 볼 수 있다.

그림 4는 NGT, LGT, 두 개의 전송 시간 그룹을 고려한 OLGT, Proposition 1의 lower bound 성능

을 비교한 것이다. 모든 데이터 사용자를 동시에 전송하는 경우는 Proposition 1에 의하여 명백하게 lower bound 성능이 가장 좋음을 알 수 있다. 그림 3과 그림 4는 4가지 방식 모두 음성 사용자와 데이터 사용자가 증가할수록 평균 전송 지연도 증가하고 있음을 보여주고 있는데, 데이터 사용자의 수가 증가할수록 상대적으로 LGT 방식 및 OLGT 방식의 성능이 그림 5는 전송 제한률을 386kbps로 제한할 경우에 최적의 전송 시간 그룹을 고려한 OLGT 방식과 SPTF 방식의 성능을 비교한 것이다. 혼합 사용자의 수의 관계없이 항상 OLGT의 성능이 좋은 것을 알 수 있다. 표 1은 제한 전송률이 주어졌을 때 최적의 전송 시간 그룹의 수 및 OLGT와 SPTF의 성능을 비교한 것이다. 즉, 전송률이 제한될 경우에는 본 논문에서 제안된 스케줄링 방식이 SPTF 방식보다 성능이 좋은 영역이 존재함을 볼 수 있다.

표 1. 최적의 전송 시간 그룹핑 및 성능 비교그림

R_M	N_v	N_d	N_{OLGT}	$N = \lceil N_d / N_{OLGT} \rceil$	Average Trans. Delay 성능비교
64 kbps	1	10	10	1	$E_{OLGT}(D) < E_{SPTF}(D)$
		25-300	12	3-25	
	50	10-300	10	1-25	$E_{SPTF}(D) < E_{OLGT}(D)$
		10-300	8	2-38	
	200	10	3	4-100	$E_{SPTF}(D) < E_{OLGT}(D)$
25-300		$E_{OLGT}(D) < E_{SPTF}(D)$			
144 kbps	1	10-300	6	2-50	$E_{OLGT}(D) < E_{SPTF}(D)$
		50	10-300	5	
	100	10-300	4	3-75	$E_{OLGT}(D) \approx E_{SPTF}(D)$
		200	10-300	1	
384 kbps	10	10-300	4	3-75	$E_{OLGT}(D) < E_{SPTF}(D)$
		50	10	3	4-100
	25-300		$E_{OLGT}(D) < E_{SPTF}(D)$		
	100	10-300	2	5-300	$E_{OLGT}(D) < E_{SPTF}(D)$
		200	10-300	1	10-300(SPTF)



3. 여러가지 스케줄링 방식의 성능 비교 ($R_i < \infty$)

그림6은 전송률 제한이 있는 경우 평균 전송 지연 및 패킷 전송이 끝나는 시점들의 분산을 나타낸 것이다. SPTF 방식은 음성 사용자와 데이터 사용자가 증가하면 분산이 급격하게 증가하고 있으며, OLGT 방식은 오히려 분산이 감소함을 볼 수 있다. 평균 전송 지연의 차이가 크면 분산의 차이도 커짐을 볼 수 있으며, 그러므로 분산이 작아지면 평균 전송 지연 역시 작아질 것이다. 이는 제안된 OLGT 방식이 Proposition 1,2의 관점에서 SPTF 방식보다 최적의 스케줄링 방식에 좀 더 근접함을 암시한다.

VI. 결론

본 논문에서는 버스트 교환 방식을 기본으로 하는 DS/CDMA 시스템의 통합 서비스에서 데이터 사용자를 특정한 전송 우선 순위를 주는 메트릭에 의한 그룹 전송 기법과 평균 전송 지연이 최소화 되도록 전송률을 할당하는 스케줄링 하는 방식을 다루었다. 단위 서버 모델에 기반을 둔 최우선 순위 에 있는 한명의 사용자에게만 전송률을 할당하는 SPTF 방식에 비하여 제안된 방식은 다중 서버 모델을 고려한 방식으로서, 모든 데이터 사용자를 한 개의 그룹으로 간주하여 동시에 모두 전송하는 기존의 LGT 방식 [9]과 달리 데이터 사용자를 여러

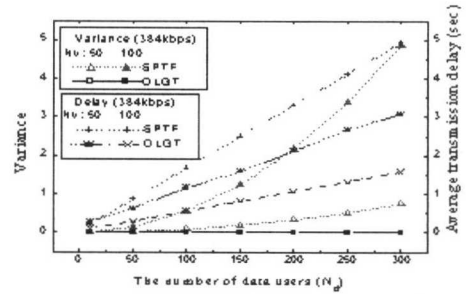


그림 6. 평균 전송 지연과 전송 완료 시간의 분산 비교

개의 전송 그룹으로 구분하여 각 그룹에 속한 버스트들을 특정한 순서대로 전송하고 각각의 전송 그룹 또한 순서대로 전송하는 OLGT 방식을 제안하였다. LGT 방식에서는 평균 전송 지연이 최소화 되도록 모든 버스트를 전송 우선 순위 에 따른 두개의 서브 그룹으로 구분하여 각각의 그룹에 최적의 전송률을 할당하는 방식을 다루는 반면, OLGT 방식에서는 데이터 사용자수가 많은 경우 데이터 사용자를 전송 우선 순위 에 따른 여러 개의 전송 그룹으로 나누어 각각의 그룹에 LGT 방식을 적용함으로써 평균 전송 성능을 추가로 향상시킬 수 있었다. 분석에서 본 바와 같이 모든 데이터 사용자에게 단일의 전송률을 할당하는 NGT 방식보다는 성능이 보다 향상되었고, SPTF 방식에 비하여서도 자원 효율성뿐만 아니라 평균 전송 지연 관점에서도 성능이 향상된 것을 볼 수 있다. LGT 방식과 마찬가지로 OLGT 방식도 최적의 스케줄링 방식은 아니지만 지원 가능한 최대 전송률의 제한이 없는 경우에는 OLGT 방식에서 전송 그룹의 수를 증가시킬수록 성능이 현저하게 향상됨을 볼 수 있다. 또한, 최적의 전송 그룹을 고려함으로써 최적해에 수렴함을 간접적으로 유추할 수 있고, 전송률의 제한이 있는 경우에는 최적의 전송 그룹을 결정함으로써 SPTF보다 우수한 성능을 나타낼 수 있다.

제안된 OLGT 방식은 데이터 사용자의 수가 많은 경우에 최대 허용 전송률의 크기에 관계없이 사용자들간의 공정성을 보장하고 효율적인 스케줄링 방식임이 판명되었다. 특히, 최적의 스케줄링 방식과 비교해서 복잡도가 매우 낮으면서도 평균 데이터 전송 지연 시간을 최소화할 수 있는 효율적인 접근 방식으로 간주된다. 한편, 우선권 순위 메트릭으로 패킷의 길이만을 고려하는 등과 같이 단순화된 부분도 있지만, 이는 성능 요구 사항에 따라 다양한 우선권 순위 메트릭을 도입함으로써 보다 일

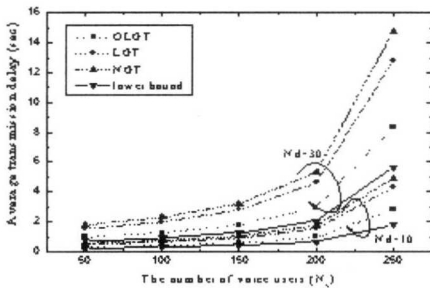


그림 4. 혼합 트래픽일 때의 성능 비교 ($R_t ∞)$

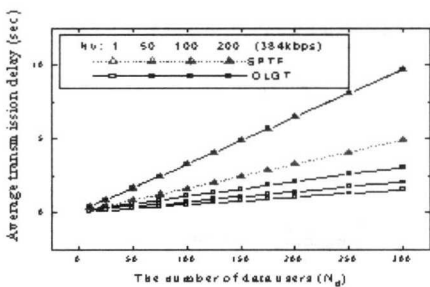


그림 5. 혼합 트래픽일 때의 성능 비교 ($R_t $\le R_M = 384$kbps)$

반화된 적용이 가능할 수 있다. 예를 들어, 전송 대기 시간, 지연 요구 사항, 채널의 상태 등을 등과 같은 요인을 우선권 순위 메트릭에 반영할 수 있을 것이다.

참 고 문 헌

[1] M. O. Sunay, Ş. Tekinay, and Ş.Z. Özer, Efficient Allocation of Radio Resources for CDMA Based Wireless Packet Data Systems, Proceedings of *IEEE Globecom'99*, vol.1, pp.638-643, Dec. 1999.

[2] C.2005-A, Upper Layer (Layer 3) Signaling Standard for cdma2000 Spread Spectrum Systems, March 2000.

[3] M. Andrews, K. Kumaran, K. Ramanan, A. Stolyar, and P. Whiting, Data Rate Scheduling Algorithms and Capacity Estimates for the CDMA Forward Link, Bell Labs Technical Journal, September 13, 1999.

[4] A. C. Varsou, H. C. Huang, and L. Mailaender, Rate Scheduling for CDMA Downlink Mixed Traffic Networks, Proceedings of WCNC 2000.

[5] A. C. Varsou and H. V. Poor, HOLPRO: A New Rate Scheduling Algorithm for the Downlink of CDMA Networks, Proceedings of Vehicular Technology Conference (Fall), September 2000.

[6] H. Kang and K. Kim, Throughput enhancement scheme for integrated voice/data DS-CDMA system: Rated-based grouping transmission, Electronics Letters, vol.35, no.17,pp.1437-1438, August 1999.

[7] S.-J. Oh and K. M. Wasserman, Optimality of greedy power control and variable spreading gain in multi-class CDMA mobile networks, in Proc. ACM/IEEE MobiCom'99, 1999, pp. 102-112.

[8] R. Jantti and S.-L. Kim, Transmission Rate Scheduling for the Non-Real-Time Data in a Cellular CDMA System, IEEE Commun. Letters, vol. 5, no. 5, May 2001.

[9] M. Kim, C. G. Kang, and R. R. Rao, Group-wise Transmission Rate Scheduling Scheme for Integrated Voice/Date Service in Burst-

Switching DS/CDMA System, IEICE Trans. Commun., E85-B, no. 8, August 2002.

[10] T. Ito, S. Sampei, and N. Morinaga, A Wireless Packet Transmission with Adaptive Processing Gain and Transmitter Power Control Scheme for Circuit-Switched and Packet-Switched Modes Integrated DS/CDMA System, IEEE Vehicle Technology Conference, pp. 2039-2043, 1999.

[11] M. Pinedo, Scheduling: Theory, Algorithms, and Systems, Prentice Hall, 1995.

[12] W. Feller, An introduction to probability theory and its applications, vol. II, 2nd ed., John Wiley & Sons, Inc, 1971.

부 록

부록 1. Proposition 1의 증명

N 명의 데이터 사용자에 대해 동시에 전송률을 할당하면서 (PO)의 최소값을 찾기 위하여 다음과 같은 전력 제약식을 만족하는 P_i 를 찾아야 한다.

$$\Delta_v + P_1 + \dots + P_N = 1, \quad \tilde{R}_i^{(d)} = P_i W / (\eta_d (1 - P_i)) \quad (A.1)$$

구한 $\tilde{R}_i^{(d)}$ 를 이용한 평균 패킷 전송 지연 D 는 다음과 같이 주어진다.

$$D = \frac{1}{N} \sum_{i=1}^N \frac{L_i}{R_i} = \frac{1}{N} \sum_{i=1}^N \frac{\eta_d L_i (1 - P_i)}{P_i W}$$

Lagrange multiplier 기법을 적용하면, Lagrange 승수 λ 에 대하여 다음과 같은 방정식이 성립하므로

$$\frac{\partial D}{\partial P_i} = -\frac{\eta_d}{NW} \frac{L_i}{P_i^2} = \lambda, \quad i = 1, \dots, N$$

이를 통해 다음의 관계식을 얻게 된다.

$$L_i / P_i^2 = \dots = L_N / P_N^2$$

이 결과와 (A.1)를 이용하여 전력 P_i , 전송률 $\tilde{R}_i^{(d)}$, 그리고 평균 전송 지연을 구하면 각각 다음과 같다.

$$P_i = (1 - \Delta_v) \sqrt{L_i} / \sum_{j=1}^N \sqrt{L_j},$$

$$\tilde{R}_i^{(d)} = \frac{(1-\Delta_i)W\sqrt{L_i}}{\eta_d \left(\sum_{i=1}^N \sqrt{L_i} - (1-\Delta_i)\sqrt{L_i} \right)}$$

$$E_{th}(D) = \frac{\eta_d}{NW} \left(\frac{(\sum_{i=1}^N \sqrt{L_i})^2}{1-\Delta_i} - \sum_{i=1}^N L_i \right)$$

부록 2. Proposition 2의 증명

Proposition 2를 증명하기 위하여 다음 Lemma 들을 먼저 증명한다.

Lemma 1. (P1)을 최소화하기 위하여는 전력 제약 식에서 등식을 만족하는 전송률 벡터 $\mathbf{R}^1 = (R_1^1, \dots, R_N^1)$ 가 사용되어야 한다.

증명) 전력을 모두 사용하지 않고 (P1)이 최소가 되었다고 하고 그 때의 전송률 벡터를 $\mathbf{R}^1 = (R_1^1, \dots, R_N^1)$ 라 하자. (P1)의 최소값이 길이가 L_k 인 패킷의 전송 시간, 즉 $\min_{R \in F_0} \max_{1 \leq i \leq N} (L_i/R_i^1) = L_k/R_k^1$ 이라고 하면 남은 전력이 있으므로, 적당한 $\delta > 0$ 에 대하여 $L_k/(R_k^1 + \delta) < L_k/R_k^1$ 이 성립한다. 즉 $\mathbf{R}^2 = (R_1^1, \dots, R_k^1 + \delta, \dots, R_N^1)$ 이 (P1)을 최소화하는 벡터이다. 이는 가정에 모순이므로 전력은 모두 사용되어야 한다.

Lemma 2. $\mathbf{R}^1, \mathbf{R}^2 \in F_0$ 가 전력을 모두 사용하는 전송률 벡터로서 $\max_{1 \leq i \leq N} (L_i/R_i^1) \leq \max_{1 \leq i \leq N} (L_i/R_i^2)$ 을 만족한다고 하면 $Var_{\mathbf{R}^1}(L_i/R_i^1) \leq Var_{\mathbf{R}^2}(L_i/R_i^2)$ 가 성립하고 역도 성립한다.

증명) (P1)을 최소화 하려면 긴 패킷에 높은 전송률을 할당하여야 함은 명백하다. L_k 가 가장 긴 패킷의 길이라고 하면 일반성을 잃지 않고 전송률 벡터 $\mathbf{R}^1, \mathbf{R}^2$ 모두에서 R_1^1, R_2^2 가 각각 패킷에 할당된 가장 큰 전송률이라고 하자. 가정 $\max_{1 \leq i \leq N} (L_i/R_i^1) \leq \max_{1 \leq i \leq N} (L_i/R_i^2)$ 에 의하여 $R_k^1 \geq R_k^2$ 이 성립하고 $\mathbf{R}^1, \mathbf{R}^2$ 모두 모든 전력을 다 사용하고 있으므로 $R_j^1 \leq R_j^2$ 인 j 가 존재한다. 일반성을 잃지 않고 벡터 $\mathbf{R}^1, \mathbf{R}^2$ 에서 j, k 성분을 제외한 나머지 성분은 모두 같다고 하자. 그러면 $L_j/R_j^1 \leq L_j/R_j^2 \leq L_k/R_k^1 \leq L_k/R_k^2$

를 얻고 그러므로 $Var_{\mathbf{R}^1}(L_i/R_i^1) \leq Var_{\mathbf{R}^2}(L_i/R_i^2)$ 이 성립한다.

역으로 $Var_{\mathbf{R}^1}(L_i/R_i^1) \leq Var_{\mathbf{R}^2}(L_i/R_i^2)$ 이 성립한다고 하면 전송률 벡터 \mathbf{R}^1 에 의하여 전송되는 패킷의 전송 완료 시간의 차이가 작으므로 패킷이 길이가 긴 것에 더 많은 전송률을 할당한 경우이다. 그러므로 전송 시간의 분산이 크게 되도록 하는 전송률 벡터 \mathbf{R}^2 에 의해 전송된 경우의 가장 긴 전송 완료 시간은 \mathbf{R}^1 에 의한 것보다 크다. 그러므로 $\max_{1 \leq i \leq N} (L_i/R_i^1) \leq \max_{1 \leq i \leq N} (L_i/R_i^2)$ 을 얻는다.

Lemma 2로부터 (P1)이 최소가 되도록 하는 전송률 벡터는 각 패킷의 전송이 끝나는 시점의 분산이 최소가 되게 하는 것임을 알 수 있다.

위의 사실을 이용하여 각 패킷의 전송이 끝나는 시점의 분산이 최소가 되도록 하는 관계식을 유도하면 된다, 즉

$$\min_{R \in F_0} Var(L_i/\bar{R}_i^{(d)}) = \min_{R \in F_0} E(L_i/\bar{R}_i^{(d)} - E(L_i/\bar{R}_i^{(d)}))^2 = 0$$

위의 관계식은 모든 i 에 대하여 $L_i/\bar{R}_i^{(d)} = E(L_i/\bar{R}_i^{(d)})$ 이 성립하는 경우, 다음을 만족하는

$$L_1/\bar{R}_1^{(d)} = L_2/\bar{R}_2^{(d)} = \dots = L_N/\bar{R}_N^{(d)} = E(L_i/\bar{R}_i^{(d)})$$

전송률 벡터 $\bar{\mathbf{R}} = (\bar{R}_1^{(d)}, \dots, \bar{R}_N^{(d)})$ 에 대하여 (P1)은 최소값을 갖는다.

부록 3. (Q0)와 (Q1)의 동일성(equivalence) 증명

전송구간동안 음성 사용자의 수는 일정하다고 가정하므로 남은 전력은 항상 일정하다고 가정하고 사용자간의 간섭이 작을수록 성능이 좋을 것이므로 가능하면 동시에 전송하고 있는 패킷의 수는 적을수록 효율적이라고 하자. 최대 가능 전송률의 제한이 있는 경우와 제한이 없는 경우로 나누어서 살펴본다.

Case 1: $R_i < \infty$

전송 제한률이 있는 경우, 즉 $R_i < \infty$ 인 경우는 SPTF 방식이 가장 좋은 성능을 나타낼 것이므로

(Q0)에서의 $\mathbf{t} = (t_1, \dots, t_{N_d})$ 는 $t_1 < \dots < t_{N_d}$ 을 만족한다.
 (Q1)에서의 $\mathbf{t}' = (t'_{11}, \dots, t'_{N_d,1})$ 는 $t'_{11} = t_1, \dots, t'_{N_d,1} = t_{N_d}$ 로 놓고 그 때의 $R_i = R_{j_i}$ 는 남은 전력을 전부 할당하면 되며 또한 (\mathbf{n}, r) 은 $\underbrace{(1, \dots, 1)}_{N_d}$ 가 된다.

정을 역으로 고려하면 (Q0)에서의 벡터 (\mathbf{t}, \mathbf{R}) 를 다음과 같이 구할 수 있다.

$$t_{(j-1)N+i} = t_{j_i}, R_{(j-1)N+i} = R_{j_i} = R_M, \forall j, i$$

이 벡터 (\mathbf{t}, \mathbf{R}) 는 F_1 의 원소이며, 이 벡터에 대하여 (Q0)는 최소값을 갖는다.

Case 2: $R_i \leq R_M$

전송 제한률이 있는 경우, 즉 $R_i \leq R_M, R_M < \infty$, 인 경우는 스케줄링이 시작된 순간에 동시에 R_M 으로 전송할 수 있는 패킷의 수 N 를 전력 제한식을 이용하여 먼저 결정한다.

먼저 (Q0)를 최소화하는 벡터에서 (Q1)을 최소화하는 벡터로의 대응을 살펴보자. (Q0)에서의 벡터 (\mathbf{t}, \mathbf{R}) 는 다음과 같이 표시된다.

$$t_1 = \dots = t_N = 0, t_{N+1} = \min_{1 \leq i \leq N} \{L_i / R_M\}$$

$$t_{kN+j} = \min_{1 \leq i \leq N} \{L_i / R_M : 0 \neq L_i / R_M \neq t_m, m = 1, \dots, kN + j - 1\},$$

모든 $i = 1, \dots, N_d$ 에 대하여 $R_i = R_M$ 이다.

이에 대응하는 (Q1)의 벡터 $(\mathbf{t}', \mathbf{R}, (\mathbf{n}, r))$ 는 다음과 같이 구성된다.

$$\mathbf{t}' = ((t'_{11} = t_1, \dots, t'_{1N} = t_N), \dots, (t'_{n1} = t_{(n-1)N+1}, \dots, t'_{nN} = t_{(n-1)N+N}),$$

$$(t'_{(n+1)1} = t_{nN+1}, \dots, t'_{(n+1)r} = t_{nN+r}))$$

$$R_i = R_M, (\mathbf{n}, r) = \underbrace{(N, \dots, N, r)}_n$$

즉 (Q0)를 최소화하는 벡터에 대응하는 (Q1)을 최소화 하는 벡터를 F_2 에서 구할 수 있다. 이 벡터에 의하여 (Q1)이 최소화 되는 것은 명백하다.

다음으로 (Q1)을 최소화하는 벡터에서 (Q0)를 최소화하는 벡터로의 대응을 살펴보자. 마찬가지로 (Q1)을 최소화하는 벡터 $(\mathbf{t}', \mathbf{R}, (\mathbf{n}, r))$ 가 다음과 같이 주어져 있다고 하자.

$$(\mathbf{t}', \mathbf{R}, (\mathbf{n}, r)) = ((t'_{11}, \dots, t'_{1N}, \dots, t'_{n1}, \dots, t'_{nN}, t'_{(n+1)1}, \dots, t'_{(n+1)r}),$$

$$R_M = R_M, \forall j, i, (N, \dots, N, r))$$

여기서 N 은 위의 경우와 같은 스케줄링이 시작된 순간에 동시에 R_M 으로 전송할 수 있는 패킷의 수이다.

각각의 그룹에 속한 각각의 패킷의 전송 완료 시간을 다르므로 첫번째 그룹을 제외하고는 나머지 그룹에 속한 패킷들은 동시 전송이 아니다. (1)의 과